Image Resolution of a Holographic System
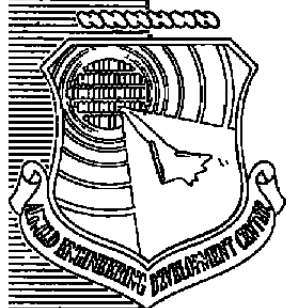
R. W. Menzel
ARO, Inc.

July 1981

Final Report for Period October 1, 1979 — July 30, 1980

ARNOLD ENGINEERING DEVELOPMENT CENTER
ARNOLD AIR FORCE STATION, TENNESSEE
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.
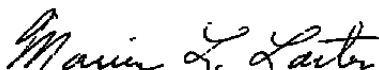
APPROVAL STATEMENT

This report has been reviewed and approved.

MARSHALL K. KINGERY
Directorate of Technology
Deputy for Operations

Approved for publication:

FOR THE COMMANDER

MARION L. LASTER
Director of Technology
Deputy for Operations

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1 REPORT NUMBER AEDC-TR-80-41	2 GOVT ACCESSION NO.	3 RECIPIENT'S CATALOG NUMBER
4 TITLE (and Subtitle) IMAGE RESOLUTION OF A HOLOGRAPHIC SYSTEM		5 TYPE OF REPORT & PERIOD COVERED Final Report–October 1, 1979 to July 30, 1980
		6 PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) R. W. Menzel, ARO, Inc., a Sverdrup Corporation Company		8 CONTRACT OR GRANT NUMBER(s)
9 PERFORMING ORGANIZATION NAME AND ADDRESS Arnold Engineering Development Center/DOT Air Force Systems Command Arnold Air Force Station, Tennessee 37389		10 PROGRAM ELEMENT PROJECT, TASK AREA & WORK UNIT NUMBERS Program Element 65807F
11 CONTROLLING OFFICE NAME AND ADDRESS Arnold Engineering Development Center/DOS Air Force Systems Command Arnold Air Force Station, Tennessee 37389		12 REPORT DATE July 1981
		13. NUMBER OF PAGES 29
14 MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15 SECURITY CLASS. (of this report) UNCLASSIFIED
		15a DECLASSIFICATION/DOWNGRADING SCHEDULE N/A

16 DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17 DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18 SUPPLEMENTARY NOTES

Available in Defense Technical Information Center (DTIC)

19 KEY WORDS (Continue on reverse side if necessary and identify by block number)

functions (mathematics)   images
linear systems   resolution
transfer functions
holography

20 ABSTRACT (Continue on reverse side if necessary and identify by block number)

The line spread function of the real image reconstruction of
a holographic system is experimentally obtained for two cases that
differ in the apertured size of the hologram. The theoretical
derivations of the holographic reconstruction of point and line
spread functions are outlined, and the multiparametric nature of
these functions is indicated. Linear systems analysis is used to
detail an experimental technique for obtaining the line spread

DD FORM 1473   EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73

20.   ABSTRACT (Continued)

functions.   In terms of the resolution each implies, the experimental results and theoretical model results agree within a factor of two.

.

AFSC
Arnold AFS Tenn

PREFACE

The work reported herein was conducted by the Arnold Engineering Development Center (AEDC), Air Force Systems Command (AFSC). The results of the research were obtained by ARO, Inc., AEDC Group (a Sverdrup Corporation Company), operating contractor for the AEDC, AFSC, Arnold Air Force Station, Tennessee, under ARO Project Number P32L-02. The Air Force project manager was Mr. Marshall Kingery. The manuscript was submitted for publication on August 13, 1980.

Mr. R. W. Menzel is currently employed by Calspan Field Services, Inc., AEDC Division.

CONTENTS

ILLUSTRATIONS

1.0 INTRODUCTION

The usefulness of holography as a data-gathering tool is determined primarily by the accuracy with which data can be reconstructed from the hologram. Various reconstruction characteristics have been studied extensively. Some of these aspects are the image intensity distribution (Refs. 1 and 2), resolution (Refs. 3 and 4), aberration (Refs. 5 and 6), focal tolerance (Refs. 7 and 8), and film effects (Refs. 9 and 10). Because all of these affect the clarity and acuity of the images and, hence, their measurement, it is not surprising that a single, direct, universally applicable definition of information resolution within the images is not available. The ubiquitous Rayleigh resolution, though simple in concept and application, does not correlate with experimentation any better than in that it implies an upper boundary on the resolution. The Sparrow resolution requirement correlates more highly with observed results; however, its physical interpretation is less clear. Attempts to define simple resolution seem to be thwarted by the facts that line images are more easily recognized than circular images of comparable diffracting energies and that line images with line widths considerably narrower than the diameters of the smallest observable circular images are easily detected.

More meaningful resolution information can be obtained by applying transfer function and linear systems theory to optical systems. This has also been applied to holographic image analysis (Refs. 11 and 12). The power of the technique resides in the superposition principle of linear systems (i.e., the response of the system to any one excitation is not affected by responses to other excitations). From this it follows that if the transform characteristics of the system for an impulse function are known, then the output response to a complicated input function is the superposition integral (the convolution) of the impulse function over the input function domain. Likewise, the Fourier transform of the output function is the product of the Fourier transform of the input function and the transfer function. In terms of optics, the impulse function is a point radiation source. By obtaining the transform of the optical system to the point radiation source, which is the point spread function, the essential features of the subsequent image characteristics can be found. A related and more easily experimentally obtainable transfer function is the line spread function resulting from a line radiating source.

Holographic systems are coherent imaging systems and, as such, are linear in the complex amplitude of their transmission and nonlinear in the intensity distribution. The conventional optical transfer and modulation transfer functions of the linear systems for incoherent imaging characteristics are not applicable to coherent images. The transfer function of coherent images cannot be determined, in general, from a Fourier transform of the intensity of the spread function. However, for the in-focus image of an idealized holographic reconstruction, as described below, the intensity distribution is the same as that

for a conventional lens imaging system. This suggests that when only the in-focus image intensity of the reconstruction of a point or line source is considered, a point spread or line spread relationship is obtained that may be used to infer the resolution of the holographic system. From this view point, the linear systems theory is applied in correlating the intensity distribution of a known point or line radiation source with the intensity distribution that results from the reconstructed image of that source. The transfer function derived in this linear systems approach is not one whose convolution with an arbitrary input function will determine the output function. Rather, it describes the deterioration in the image of the original radiation source.

Thus, resolution of a holographic system is inferred from the spread function derived by using a linear systems approach. This report describes the process of arriving at the results. The major sections of the report are (1) an outline of the theoretical point and line spread function of a holographic system, (2) a discussion of linear systems theory to allow a thorough description of a method for obtaining the line spread functions, (3) the experimentally obtained line spread functions for a holographic system, and (4) a comparison of theory, experiment, and observed resolution of holographic images.

2.0 HOLOGRAPHIC IMAGE OF A POINT AND A LINE SOURCE

A mathematical model of the reconstructed image of a point source of radiation can be formulated through consideration of the holographic process in its component parts. The geometric parameters of the system are shown in Fig. 1. The (x_i, y_i) planes with $i = 1, 2,$ and 3 are, respectively, the point source plane, the hologram plane (for both recording and reconstruction), and the reconstructed image plane. For simplicity, it is assumed that the reference beam is collimated and that the reconstruction beam is the same except that its angle is $-\psi$. The point source is assumed to be at the origin of the (x_1, y_1) plane.
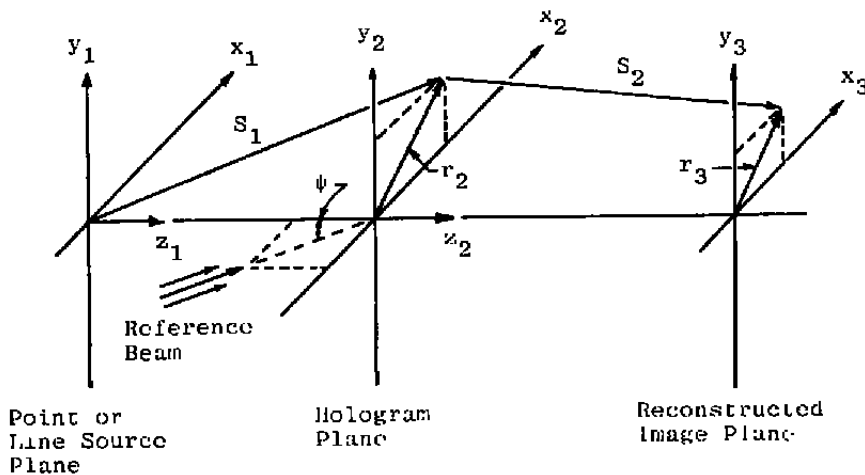

Figure 1. Definition of the holocamera parameters.

6

The point radiation source is described by Green's function $(A/s_1) \exp (iks_1)$, where $k = 2\pi/\lambda$ and λ is the wavelength. The holographic recording is the interference between the Green's function and the reference wave. With the binomial expansion $s_1 \cong z_1 + 1/2(r_2^2/z_1)$ and the assumption that the point source and reference wave have the same amplitude, A, then the intensity in the hologram plane is given by

$$I_H = A^2 \left[1 + (1/z_1)^2 + (2/z_1) \cos (kz_1 + kr_1^2/2z_1 - kx_2 \sin \psi) \right] \tag{1}$$

When $\psi = 0$, the intensity distribution is similar to a Fresnel zone plate.

The reconstructed amplitude field, U_R, is the Fresnel diffraction obtained from the hologram,

$$U_R = -\frac{i}{\lambda} \iint_{\text{aperture}} U_H \left(\frac{e^{iks_2}}{s_2} \right) dx_2 \, dy_2 \tag{2}$$

where U_H is the amplitude field transmitted by the illuminated hologram. U_H is given by the product of the incident illuminating field and the complex transmittance of the hologram. If phase shifts caused by the film emulsion thickness variations are neglected, then the hologram's complex amplitude transmittance can be approximated by a linear mapping of the complex amplitude during exposure, provided that the exposure level is properly chosen (Ref. 11). Using the binomial expansion of s_2 and the exponential expansion of the cosine in Eq. (1), and changing to cylindrical coordinates allows the reconstructed amplitude field to be written as

$$U_R = \sum_{j=1}^{3} K_j \int_0^1 J_0 (v\rho) \exp (i u_j \rho^2) \rho \, d\rho \tag{3}$$

where

$$v = kr_3 R/z_2$$

$$u_1 = x_2 k \sin \psi$$

$$u_2 = (kR^2/2) (1/z_1 + 1/z_2) - 2kx_2 \sin \psi \tag{4}$$

$$u_3 = (kR^2/2) (1/z_1 - 1/z_2)$$

R is the maximum radial dimension of the hologram (i.e., the effective aperture of the hologram). The K_j incorporates constants that do not influence the spatial distribution of intensity in the reconstructed images. The three integrals represent, successively, the reconstruction of the bias on the film, the virtual image, and the real image. Because of the angle ψ, the images are spatially separated.

7

The basic integral of the reconstruction process

$$U = \int_0^1 J_o(v\rho) \exp(i u \rho^2) \rho \, d\rho \tag{5}$$

can be solved by using Lommel functions (Ref. 13). The intensity obtained from UU* can be graphically represented by a set of isophote diagrams (Refs. 13 and 14) in which contour lines of constant intensity near the focal volume of the optical system are plotted on a (u, v) coordinate system. The importance of these isophote diagrams is that the solution they represent was obtained in analyzing the three-dimensional light distribution near the focus of a diffraction-limited lens. In other words, for the above assumptions, the intensity field in the focused region of the reconstruction of the hologram of a point source is identical in form to the intensity field near the focal volume of a diffraction-limited lens. Therefore, the focusing characteristics attributed to the lens may equally well describe the holographic reconstruction. In particular,

1. The fraction of the total energy that falls within a given area about the optical axis for any constant u is the same for point images obtained by lens projection and for holographic reconstruction.

2. The phase distribution near focus is the same in both cases (i.e., in the focal waist the phase surfaces propagate like plane waves).

The similarity of lens and holographic images shows that the physical properties of either image (under appropriate conditions, of course) do not reveal which system was used to make the image. Indeed, from this standpoint, holography is simply another imaging system; its unique properties relate to image storage capabilities but not to the image characteristics themselves.

The point spread function is the image intensity when the focusing condition $u = 0$ is satisfied. For the normalized real image, this intensity is the familiar Airy pattern associated with diffraction from an aperture of radius R,

$$I_{u=0} = \left[J_1(k r_3 R/z_2) / (k r_3 R/z_2) \right]^2 \tag{6}$$

The multiparametric nature of the point spread function is worth noting.

In principle, the point spread function is the trace that results from scanning the image with a suitably small pinhole. Experimentally, it is difficult to obtain because the energy passing through the pinhole is insufficient for accurate measurement. On the other hand, if

8

a line source of radiation is imaged, the image may be satisfactorily scanned with an elongated slit parallel to the image axis.

The derivation of the image intensity distribution which is the line spread function follows the same form as for the point spread function. For a line source of radiation along the y_1 axis, the reconstructed real image amplitude assumes the form

$$U_R = K \int_{-L}^{L} \exp(i u x_2^2) \exp(i v x_2) \, dx_2 \tag{7}$$

where now

$$v = -kx_3 L/z_2$$

and

$$u = (k L^2/2)(1/z_2 - 1/z_1) \tag{8}$$

and L is half the width of the hologram in the x_2 direction. Again, the line spread function is the image intensity when the focusing condition u = 0 is satisfied. The normalized intensity is

$$I_{u=o} = \text{sinc}^2 \left[kx_3 L/z_2 \right] \tag{9}$$

Comparing the widths of the line spread function to the point spread function, for equal arguments, at their respective first zero values shows $x_3 = 0.82 \, r_3$. Consequently, for holograms of the same size, objects with linear features are resolved with a spread function narrower than that of objects with circular symmetry. This is experimentally verifiable: Linear images are much easier to recognize in the reconstruction than circular images.

3.0 LINEAR SYSTEMS

The theoretical and experimental foundation of the application of the line spread function to the analysis of photographic systems has been developed extensively in an excellent series of articles by Lamberts et al. (Refs. 15 and 16), Perrin (Refs. 17 and 18), and Higgins et al. (Ref. 19). The Fourier transform technique of optical image evaluation (i.e., the transfer function) is described at length by Smith (Ref. 20).

The spread function and transfer function techniques are applied to a holographic system by first considering the entire holographic system — the diffraction phenomenon, the holographic recording, the reconstruction diffraction, and the recording of the

reconstructed image — as a "black box" optical device that has the useful property of transmitting an image from one place to another. The actual mechanism of the image transmission is ignored, and comparisons are made only between the "input" image and the "output" image intensity distributions. In this manner, the question of coherent or incoherent imaging does not arise. For any given input to the black box, there exists a corresponding output whose facsimile fidelity has been degraded by the transmission process. For example, the energy in the transmitted image of a step function input such as an illuminated knife edge does not have a step function distribution, but, rather, the energy "spreads" so that a portion falls into geometrical shadow of the step function.

To pursue the application of the linear systems techniques it is appropriate to note several properties of linear optical systems. An image function $I(x_i, y_i)$ is correlated with the object function $O(x_o, y_o)$ by the superposition integral (Ref. 11)

$$I(x_i, y_i) = \iint_{-\infty}^{\infty} O(x_o, y_o) S(x_i, y_i ; x_o, y_o) \, dx_o \, dy_o \tag{10}$$

where $S(x_i, y_i; x_o, y_o)$ is the system's response to an impulse function. With unit magnification, the space described by (x_o, y_o) is the same as the (x_i, y_i) space, and the object and image can be compared point for point. When S is dependent only on the coordinate differences $(x_i - x_o)$ and $(y_i - y_o)$, the linear system is space invariant. Then,

$$I(x_i, y_i) = \iint_{-\infty}^{\infty} O(x_o, y_o) S(x_i - x_o ; y_i - y_o) \, dx_o \, dy_o \tag{11}$$

which has the form of a two-dimensional convolution of the object function with the impulse response function. The condition of space invariance requires that the system operating on a moving point object result in a moving point image that does not change in its mathematical form. This requirement demands an aberration-free optical system. The convenience of the application of the convolution integral makes it desirable to use this technique with more general cases than those which are aberration free. This is obtained, in principle, when the object field domain is restricted sufficiently so that the image field is satisfactorily space invariant. Several adjacent domains may comprise the object field, and each domain may have associated with it a different impulse response.

Equation (11) may also be written as

$$I(x_i, y_i) = \iint_{-\infty}^{\infty} O(x_i - x_o ; y_i - y_o) S(x_o, y_o) \, dx_o \, dy_o \tag{12}$$

by making a simple coordinate transformation. This form of the convolution is applied in the subsequent derivations.

When the spread function of an optical system is known, the entire transformation characteristics can be derived for any image obtained with the system. For example, if the object function is a point radiation source, the resolution of the system is often defined as the halfwidth of the spread in the resulting image. If, on the other hand, the system is comprised of multiple subsystems, the spread function of the entire system is obtained by convolving the spread functions of the subsystems

$$S(x, y) = \iint S_1(x', y') \, S_2(x - x' \; ; \; y - y') \, dx' \, dy' \tag{13}$$

Conversely, the spread function of a subsystem can be found with deconvolution process. Since this is usually very difficult to perform mathematically, it is advantageous to use the Convolution Theorem by which the Fourier transform of the convolution of two functions in dimensional space is equivalent to the multiplication of the Fourier transforms of each of the two functions in spatial frequency "space." Thus, spatial deconvolution is the operational equivalent of dividing by the correct Fourier transform and finding the inverse Fourier transform of the result. For example,

$$S_1(x, y) = \mathcal{F}^{-1} \left\{ \mathcal{F}\left[S(x, y) \right] \Big/ \mathcal{F}\left[S_2(x, y) \right] \right\} \tag{14}$$

The nature of the spread function S(x, y) is further revealed by the Fourier analysis techniques. The Fourier transform of an input field O(x, y) is the complex field described by

$$\mathcal{F}\left\{ O(x, y) \right\} = \iint\limits_{-\infty}^{\infty} O(x, y) \, \exp\left[-i \, 2\pi \, (xf_x + yf_y) \right] dx \, dy \tag{15}$$

The conjugate coordinates are the space coordinates (x, y) and the spatial frequency coordinates (f_x, f_y). The components of the Fourier expansion of the Fourier transform are characteristically sinusoidal and can be expressed by

$$K(f_x, f_y) = K_0(f_x, f_y) \, \exp\left[i \, 2\pi \, (xf_x + yf_y) \right] \tag{16}$$

11

When this component wave is itself the input function of the linear system, the important result

$$
\begin{aligned}
I(x_i, y_i) = &\iint_{-\infty}^{\infty} S(x_o, y_o) K_o(f_x, f_y) \exp\left[i\, 2\pi\, f_x(x_i - x_o) + i\, 2\pi\, f_y\, (y_i - y_o) \right] dx_o\, dy_o \\
&- K_o(f_x, f_y) \exp\left[i\, 2\pi\, (x_i f_x + y_i f_y) \right] \\
&\iint_{-\infty}^{\infty} S(x_o, y_o) \exp\left[-i\, 2\pi\, (x_o f_x + y_o f_y) \right] dx_o\, dy_o \\
&- K_o(f_x, f_y) \exp\left[i\, 2\pi\, (x_i f_x + y_i f_y) \right] \mathcal{F}\left\{ S(x_o, y_o) \right\}
\end{aligned}
\tag{17}
$$

is obtained. The image field is simply the object field modified by the Fourier transform of the spread function,

$$
K_{output}(f_x, f_y) = \mathcal{F}\left\{ S(x_o, y_o) \right\} K_{input}(f_x, f_y)
\tag{18}
$$

The problem of finding $S(x, y)$ is now essentially the same as the problem of finding $\mathcal{F}\{S(x, y)\}$.

This result, carried further in the following example, leads to the experimental method of finding the spread function. Consider a one dimensional sinusoidal object field given by

$$
0(x_o, y_o) = 0(x_o) = 1 + a \cos(2\pi x_o f_x)
\tag{19}
$$

where $a \leq 1$. The image field resulting when Eq.(12) is applied is

$$
\begin{aligned}
I(x_i) = &\int_{-\infty}^{\infty} S(x_o)\, dx_o \\
&+ a \cos(2\pi x_i f_x) \int_{-\infty}^{\infty} S(x_o) \cos(2\pi x_o f_x)\, dx_o \\
&+ a \sin(2\pi x_i f_x) \int_{-\infty}^{\infty} S(x_o) \sin(2\pi x_o f_x)\, dx_o
\end{aligned}
\tag{20}
$$

The integrals each have numerical values that will be designated, respectively, M_o, M_c, and M_s. Clearly, the M_c and M_s integrals are the cosine and sine components of the Fourier transform of the spread function. These are considered to be independent variables; therefore, with substitution of the definitions

$$
M_c = M \cos\theta \quad \text{and} \quad M_s = M \sin\theta
\tag{21}
$$

Eq. (18) becomes

$$I(x_i) = M_0 + a M \cos (2\pi x_i f_x - \theta)$$ (22)

The output differs from the input by the modulation, M, and a phase angle, θ. If $S(x_0)$ is symmetrical, then the sine integral is zero, $M_s = 0$, and the phase angle is also zero. Attention can be focused on either M_c or M_s by requiring the argument $(2\pi x_i f_x)$ to cause the cosine or sine function modifying the integrals to be zero. Choosing $\sin (2\pi x_i f_x) = 0$ and assuming further that normalization is chosen so that $M_0 = 1$ and $a = 1$ produces

$$M_c = \left\{ \left[I(x_i) \right]_{max} - \left[I(x_i) \right]_{min} \right\} \Big/ \left\{ \left[I(x_i) \right]_{max} + \left[I(x_i) \right]_{min} \right\}$$ (23)

This is similar to the Michelson visibility function. $[I(X_i)]_{max}$ is obtained for $x_i = 0$ which is the same condition as requiring the impulse response function to be centrally located over a maximum amplitude value of the sinusoidal field. Similarly, $[I(X_i)]_{min}$ is obtained when a minimum of the sinusoidal field is centered over the impulse response function.

An equation structurally the same as Eq. (23) results for M_s when $\cos (2\pi x_i f_x) = 0$. Now the maxima and minima of the output function correspond to centering the impulse response function such that the slope of the input field is maximum and minimum.

The Fourier transform of the spread function (in one dimension, to follow the above example) is

$$\mathcal{F}\left\{ S(x_0) \right\} = \int_{-\infty}^{\infty} S(x_0) \exp \left[-i\, 2\pi f_x (x_i - x_0) \right] dx_0$$

$$= M_c + i M_s$$ (24)

$$= \left[M_c^2 + M_s^2 \right]^{\frac{1}{2}} \exp \left[i \tan^{-1} \left(\frac{M_s}{M_c} \right) \right]$$

By definition $[M_c^2 + M_s^2]^{1/2}$ is the modulation transfer function, MTF, while $\exp [i \tan^{-1} (M_s/M_c)]$ is the phase transfer function, PTF. The MTF represents the relative amplitude by which different spatial frequencies are transmitted through the system (Ref. 19). The PTF describes the displacement of wavefronts that sinusoidal distributions of various spatial frequencies suffer in transmission through the system (Ref. 20).

The Fourier transform of the spread function of an optical system is known as the optical transfer function, OTF, of the system. Because the Fourier transform pair (i.e., the Fourier

transform and the inverse Fourier transform) are related such that knowledge of one implies knowledge of the other, then the OTF can, like the spread function, be used to obtain all the imaging characteristics of the optical system. It is typical of good optical systems that phase variations are significant only when the MTF is small, which is usually for the higher spatial frequencies. Then the imaginary component of the OTF is quite small and the inverse Fourier transform may be satisfactorily determined using only the MTF. Because the MTF is the magnitude of the OTF, it is always positive. The spread function obtained from the inverse Fourier transform of the MTF must then be a symmetrical function. Conversely, an experimentally determined asymmetry in the spread function indicates the presence of non-negligible phase distortion.

The simplest one-dimensional object is an illuminated knife edge, mathematically a step function. Using Eq. (12) results in the following:

$$I(x_i) = \int_{-\infty}^{x} S(x_o) \, dx_o \tag{25}$$

where the integration limit, x, locates the edge of the step function. Assuming that x_o and x_i are dimensionally the same (unity magnification), then differentiation of the image

$$d\left[I(x_i)\right] / dx_i = S(x_o) \tag{26}$$

shows that the slope of the trace of an image of a knife edge is equal to the line spread function.

The relationships of the point, line, and step function flux density mapping are now examined. Because of the impractically low energy levels implied by mapping a point source with a point aperture, $S(x_o, y_o)$, cannot be directly obtained. Larger apertures require that the flux measured be the result of integrating over the receiving aperture area. In other words, $S(x_o, y_o)$ is buried in an integration mapping. Therefore, let the point-image characteristic $P(w_x)$ be defined as the flux contained within a varyingly wide slit, W_x, of infinite length about the center of a point input function. Equation (12) becomes

$$P(w_x) = \int_{-w_x/2}^{w_x/2} \int_{-\infty}^{\infty} S(x_o, y_o) \, dx_o \, dy_o \tag{27}$$

If the input field is a line source that is characterized by a constant flux along the length of the line (assume the y direction), the flux density of the line source is given by

$$L(x_i) = \int_{-\infty}^{\infty} S(x_o, y_o) \, dy_o \tag{28}$$

Similarly, when the input field is a step function, the flux density is

$$D(x_i) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} S(x_o, y_o) \, dx_o \, dy_o \tag{29}$$

The relationship between these functions is

$$P(w_x) = D(w_x/2) - D(-w_x/2) = \int_{-w_x/2}^{w_x/2} L(x) \, dx \tag{30}$$

which shows that the point image characteristics can be obtained from the point, line, variable width line, and step function sources. Specifically, the transform characteristics of a line source are given in terms of the system's response to a step function. This equation provides the basis on which the holographic line spread functions have been determined.

4.0 EXPERIMENTAL PROCEDURE

The result of Eq. (26) has only limited direct application because when a knife edge image is scanned, the edge trace incorporates the line spread function of both the optical system and of the scanning system. Attempts to separate the respective line spread functions would require the deconvolution process and its attending difficulties. However, the edge trace slope obtained when the scanning system traces an illuminated knife edge directly is, by Eq. (26), exactly the line spread function of the scanning system.

Because of the necessity for deconvolution in experimentally obtaining the line spread function of a holographic system, the transform function analysis technique [Eq. (14)] was used instead. Briefly, the method is as follows: A holographic system is used to obtain an image of a knife edge which is then recorded on film. A microdensitometer is used to obtain the edge traces of the knife edge image on the film plate. This trace includes the effects of the spread functions of both the holographic system and the microdensitometer. To eliminate the microdensitometer effects, an edge trace of the knife edge itself was also made. From the edge traces, the MTF and PTF of the holographic and the scanning systems are obtained. The MTF of the holographic system without the microdensitometer effects is obtained [analogous to Eq. (14)] by dividing the image edge trace MTF by the knife edge trace MTF. The PTF of the holographic system minus the microdensitometer contribution is obtained by subtracting the PTF of the knife edge trace from the PTF of the image trace. The derived MTF and PTF of the holographic system are then used to obtain the real and imaginary components of the OTF. These, in turn, are used to find the line spread function (LSF), which is the inverse Fourier transform of the OTF.

The method of obtaining the MTF and PTF is a graphic technique described by Scott et al. (Ref. 21), and the principles on which it is based are discussed by Shack (Ref. 22). To proceed, note that Eq. (22) states that a linear system operating on a sinusoidal field input will result in a sinusoidal field output that is changed by a factor corresponding to the Fourier transform of the impulse response function. A similar relationship exists when the input function is a periodic square wave. The square wave can be expressed in terms of its sinusoidal Fourier series components. Each component is then operated upon independent of the other components and the results are summed to give the system square wave response function. The square wave response function $\Psi(f_x)$ is, hence, expressed by

$$\Psi(f_x) = \frac{\pi}{4}\left[F(f_x) - \frac{1}{3}F(3f_x) + \frac{1}{5}F(5f_x) - \ldots\right] \tag{31}$$

where for notational convenience the Fourier transform of the spread function in one dimension given by $(x_o, x_o/3_o, \ldots, x_o/n)$ has been written as

$$\mathcal{F}\left\{S(x_o/n)\right\} = \Gamma(n\,f_x) \tag{32}$$

Finding the components of Eq. (31) is somewhat involved. Consider an optical input to the linear system [Eq. (12)] that is a square wave field with periodicity of 2α in the x_o direction. Because there exists no y_o variation, the point spread function is replaced by the line spread function

$$I(x_i) = \int_{-\infty}^{\infty} L(x_o)\,O(x_i - x_o)\,dx_o \tag{33}$$

Expanding further such that the maximum amplitude of the square wave is centered at the origin, the image at $x_i = 0$ is expressed by the expansion,

$$
\begin{aligned}
I(x_i = 0) = &\int_{-a/2}^{a/2} L(x_o)\,dx_o \\
&+ \int_{-5a/2}^{5a/2} L(x_o)\,dx_o - \int_{-3a/2}^{3a/2} L(x_o)\,dx_o \\
&+ \int_{-9a/2}^{9a/2} L(x_o)\,dx_o - \int_{-7a/2}^{7a/2} L(x_o)\,dx_o \\
&+ \ldots
\end{aligned}
\tag{34}
$$

This, together with Eq. (30), results in

$$I_{max}(x_i = 0) = \sum_{n=-N}^{N}\left\{D\left[(2n+1/2)\alpha\right] - D\left[(2n-1/2)\alpha\right]\right\} \tag{35}$$

The limit N is arbitrary. N is chosen to ensure that further terms in the series do not contribute significantly to the result. Equation (35) succeeds in describing the square wave input in terms of a step function response.

By a similar derivation, when a minimum of the square wave field is centered on the origin,

$$I_{min} (x_i = 0) \sum_{n=-N}^{N} \left\{ D\left[(2n - 1/2)\, a\right] - D\left[(2n - 3/2)\, a\right] \right\} \tag{36}$$

Analogous with Eq. (23), the cosine turn of the square wave response function is

$$\Psi_c (f_x) = (I_{max} - I_{min}/I_{max} - I_{min}) \tag{37}$$

$$= -1 + \left\{ 2/\left[D(Na) - D(-Na)\right] \right\} \sum_{n=0}^{N} (-1)^n \left\{ D\left[(n + 1/2)\, a\right] - D\left[-(n - 1/2)\, a\right] \right\}$$

Thus, by measuring the amplitude in the image response of a step function at points $x_1 = (n + 1/2)\alpha$, one obtains the square wave response $\Psi_c(f_x = 1/2\alpha)$. The full spectrum of $\Psi_c(f_x)$ is derived by changing α and repeating the measurements.

Figure 2 shows a generalized conception of a step function response and the interval spacing used to determine $\Psi_c(1/2\alpha)$. The heavy line sections correspond to the intervals in Eq. (35) and the dashed line section to those in Eq. (36). The interval $D(n\alpha/2) - D[(n - 2)\alpha/2]$ becomes zero as n becomes large, thereby providing the convergence of the series expressed in Eqs. (35), (36), and (37).

From $\Psi_c(f_x)$ the sine wave response function can be obtained by inverting Eq. (37). The inversion process is possible because the series is not actually infinite but terminates at a high frequency limit determined by the aperture of the measuring instrument. Practically, the high frequency limit is specified by the shortest interval with which the step function response curve can be measured. Beginning with the high frequency limit, the cutoff frequency, f_c, then between f_c and $f_c/3$, the cosine wave response is given by
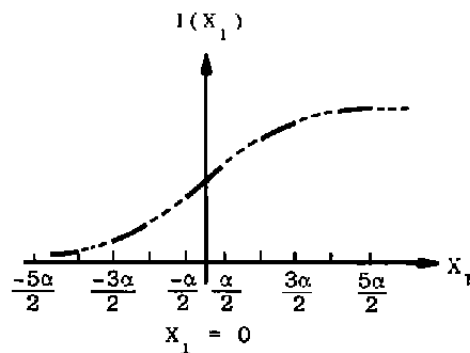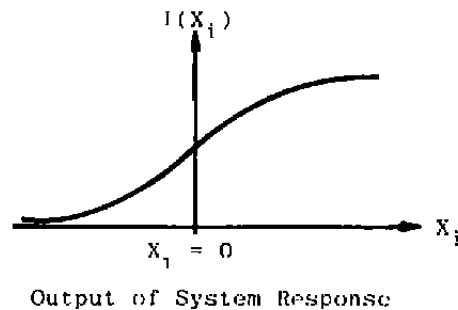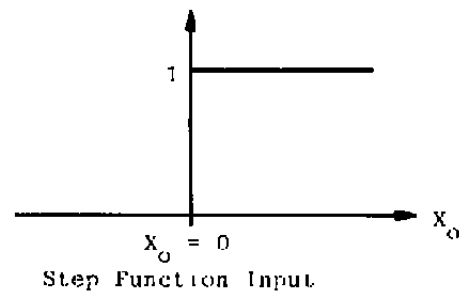
$$F_c (f_x) = \frac{\pi}{4} \Psi_c (f_x) \tag{38}$$

Between $f_c/3$ and $f_c/5$,

$$F_c (f_x) = \frac{\pi}{4} \Psi_c (f_x) + \frac{1}{3} F_c (3 f_x) \tag{39}$$

and between $f_c/5$ and $f_c/7$,

$$F_c (f_x) = \frac{\pi}{4} \Psi_c (f_x) + \frac{1}{3} F_c (3 f_x) - \frac{1}{5} F_c (5 f_x) \tag{40}$$

Step Function Input

Output of System Response

Interval Spacing for Determining $F(1/2\alpha)$

Figure 2. Conceptualization of a linear system response to a
step function input.

and so on. The low frequency limit is generally quite obvious since systems with
a d-c response become $F_c(f_x \sim 0) \to 1$, whereas those without a d-c response become
$F_c(f_x \sim 0) \to 0$.

The sine term of the square wave response function is derived analogously as

$$\Psi_s\,(f_x = 1/2\,a) = -1 + \left\{2/\left[D\,(N\alpha) - D\,(-N\alpha)\right]\right\}\;\sum_{n=0}^{N}\;(-1)^n\left[D\,(n\alpha) - D\,(-n\alpha)\right]\right\} \qquad (41)$$

The sine wave response function $F_s(f_x)$ is found in a like manner from Eqs. (38), (39), and (40). The low frequency limit is always $F_s(f_x \sim 0) \rightarrow 0$.

The modulation transfer function, MTF, is given by

$$MTF = \left[F_c^2(f_x) \cdot F_s^2(f_x) \right]^{\frac{1}{2}} \tag{42}$$

and the phase transfer function, PTF, by

$$PTF = \exp \left\{ i \tan^{-1} \left[F_s(f_x)/F_c(f_x) \right] \right\} \tag{43}$$

From $F(f_x) = F_c(f_x) + i F_s(f_x)$, the line spread function, $L(x_0)$, can be computed by numerical integration of the inverse Fourier transformation of $F(f_x)$. When symmetry can be assumed, the point spread function, $S(x_0, y_0)$, is the surface revolution of $L(x_0)$ with an axis through $x_0 = 0$.

5.0 EXPERIMENTAL RESULTS

The somewhat convoluted processes described above were applied to a holographic system by obtaining holographic images of a knife edge using the off-axis recording mode. The angle between the beams was 7 deg, and the knife edge (a razor blade) was 35 cm from the holographic recording plate (Agfa Gaevert 10E75). The object and reference beams were collimated plane waves; their intensity ratio was unity. To facilitate correct alignment of the hologram for the reconstruction, the hologram also contained the holographic recording of a pinhole which was precisely in the same plane as the knife edge. During the reconstruction, the angular alignment of the hologram was adjusted until the reconstructed pinhole image showed the least amount of aberration. This was a much more sensitive method than using the knife edge image to judge the alignment. To secure the best focus during the recording of the reconstruction, both the recording film plate and the focal position of the real image were precisely located through a microscope. The pinhole image was used to determine the focal plane. A Jarrell-Ash Console Comparator Microdensitometer was used to trace the film recording of the reconstructed knife edge image. The tracing slit was adjusted by the instrument guage to an indicated 2-μm width and 1.2 mm length; the tracing speed was 1 mm/min.

To emphasize the aperture effect in the reconstructed image, two holographic cases were investigated which differed only in hologram size. One hologram (to be referred to as the A hologram) was 2.4 cm in diameter, this value chosen as representing the nominal hologram size requirements. In anticipation of a broader spread function, the other hologram (the B hologram) was only 0.75 cm in diameter. In both cases, the geometrical image of the knife

edge was near the center of the hologram. Obviously, these parameters are such that the best resolutions possible by the holographic systems are not realized. Rather, the degrading effects of the long reconstruction distance (35 mm) and of the aperture limiting are emphasized.

Figure 3 shows the microdensitometer traces of a knife edge (a razor blade) and the reconstructed A and B hologram images of the knife edge. The traces have been normalized to remove the effects of the unexposed film emulsion absorption. Obviously, none of the traces are symmetrical, indicating the occurrence of phase distortion. Some sense of the image degradation caused by the holographic system is obtained by noting that the slope of the reconstructed image traces is not as steep as that of the knife edge trace, that the contrast between the clear and opaque fields is not as strong, and that the slope of the traces is not always positive. According to Eq. (26), this implies that regions exist in which the spread function has negative values.
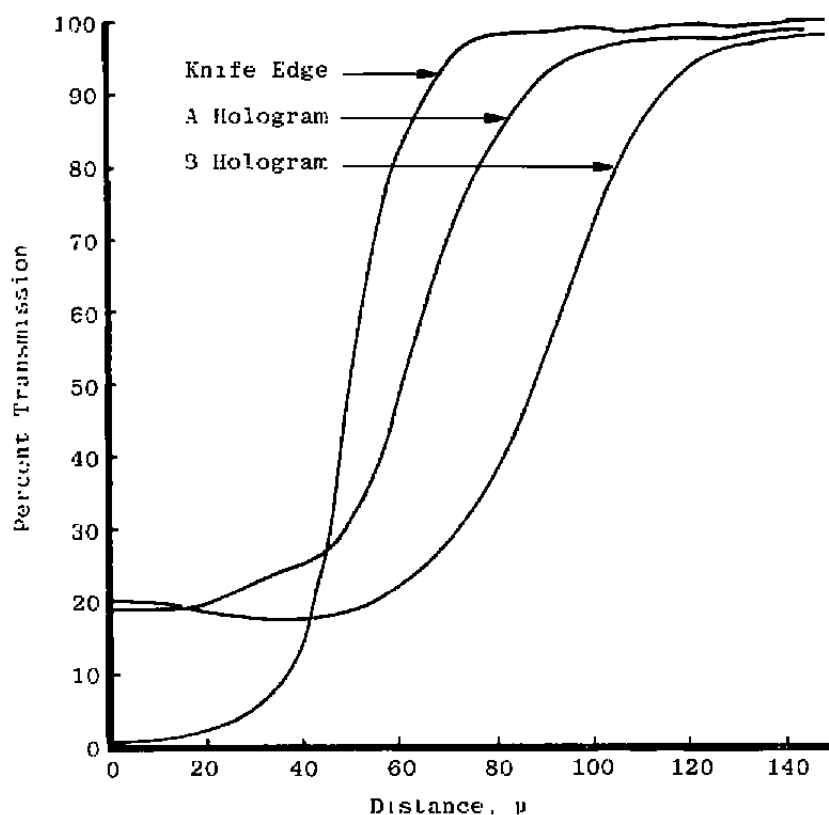

Figure 3. Microdensitometer traces.

Figure 4 shows the MTF and PTF of the microdensitometer obtained from the direct tracing of the knife edge. Some characteristics pertinent to the technique of edge trace scanning can be noted from these curves. The variation of the PTF significantly away from
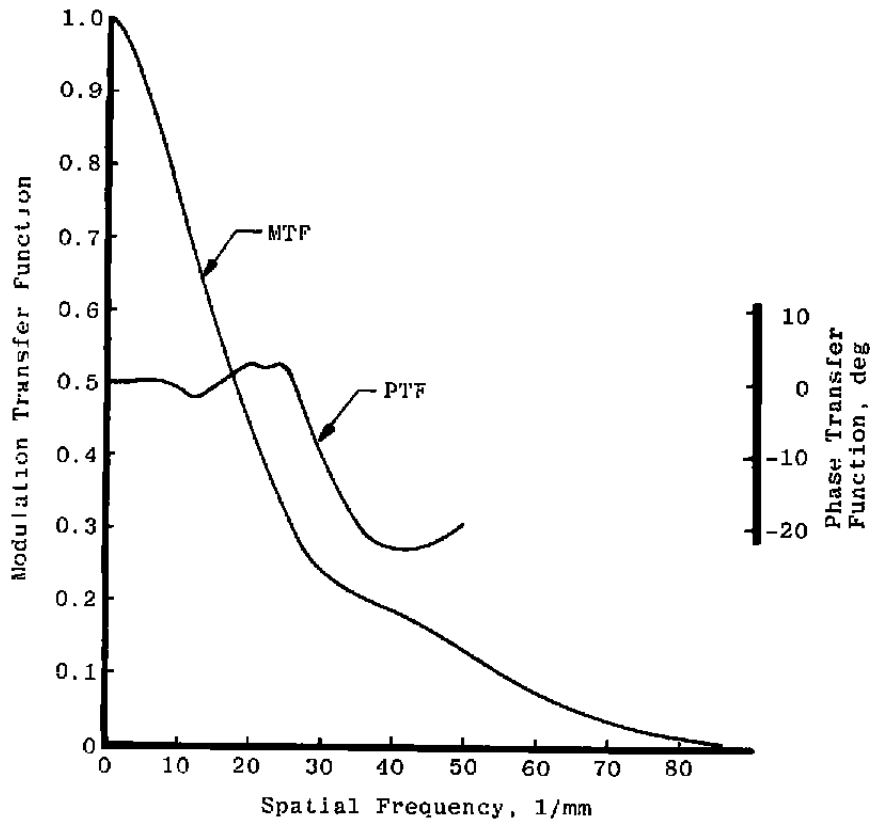
20

Figure 4. Modulation and phase transfer functions of the microdensitometer.

zero values indicates the presence of phase distortion in the microdensitometer system. The distortion may have resulted from defocusing and/or aberrations of the microdensitometer optics, angular misalignment in the slit orientations, or damaged edges of the slit. The microdensitometer was not disassembled to check for physical damage; edge integrity of the slit jaws was assumed. Geometrical alignment of the slit-trace geometry is performed visually on the Jarrell-Ash microdensitometer, and an orientation of better than within one degree is easily attained. The probable cause of the PTF variations is a slight defocusing of the microdensitometer optics.

Because the MTF decreases rather gradually, a clear cutoff frequency is not apparent. The transfer function of a microdensitometer slit is given approximately (Ref. 21) by sinc (πwf,), where w is the effective width of the slit. The use of an estimated cutoff spatial frequency of 85 to 100 (1/mm) suggests an effective slit width of 10 μm to 12 μm. The disparity between this width and the instrument setting of 2 μm implies a large discrepancy. A similar problem was addressed by Jones and Coughlin (Ref. 23), who found that correction for the microdensitometer scanning of photographic images by dividing the MTF of the photographic image by the microdensitometer MTF resulted in an "overcorrection"

21

at the higher spatial frequencies. They developed the more exact but also more difficult technique of incorporating a microdensitometer correction function into the data reduction analysis. As part of their technique, the experimentally determined microdensitometer spatial frequency is truncated at a value corresponding to its 0.1 MTF value. Smaller values are not considered sufficiently reliable. Truncation of the MTF at the 0.1 value does not, of course, allow a direct comparison of the effective slit width indicated by the spatial frequency cutoff with the actual slit width used. Nevertheless, the overcorrection effect satisfactorily explains the apparent discrepancy noted above.

Because of the large number of arithmetic computations involved, the use of a microdensitometer correction function is not practical without extensive computer programming. In the present study, time limitations prevented this; however, the work of Jones and Coughlin shows that including the microdensitometer correction function improves the details but does not significantly change the overall characteristics of the results.

Figure 5 reinforces this observation. The two sets of data are the line spread functions of the microdensitometer derived from the transfer function (the solid lines) and those derived
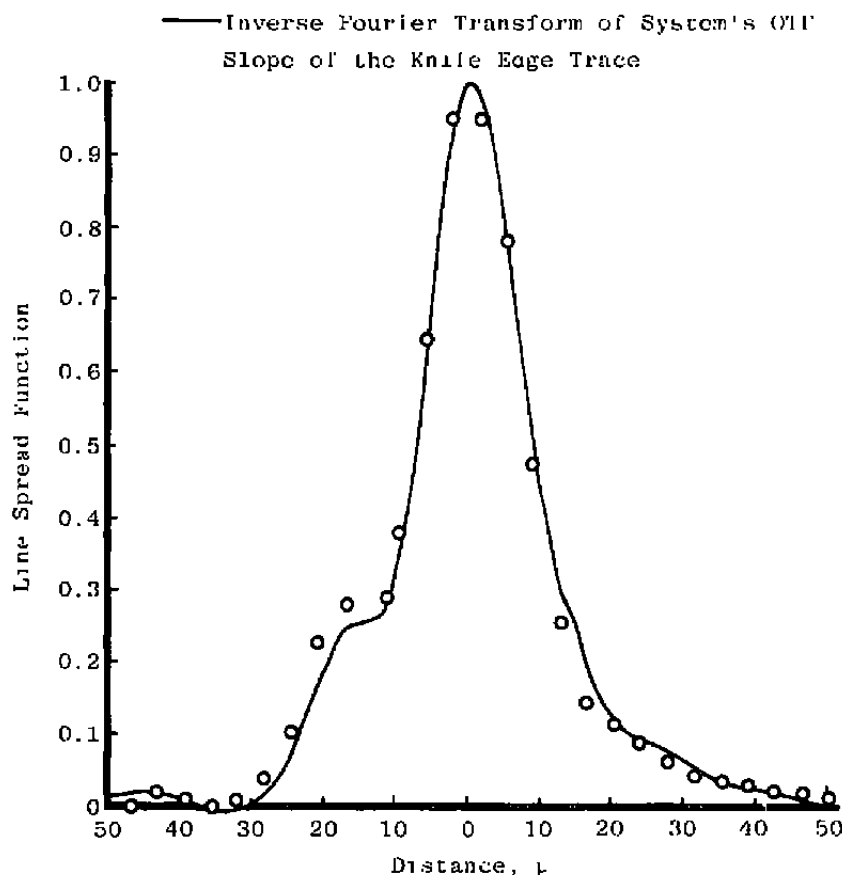
Figure 5. Normalized line spread function of the microdensitometer.

from a direct calculation of the slope of the knife edge trace (the circles), normalized for unity at the peak of the curve. The high degree of correlation justifies using the transfer function method of obtaining the line spread function, and the correlation also increases confidence in the results obtained for the holographic systems.

Figures 6 and 7 are the MTF and PTF of the A and B holographic systems, respectively. The cutoff frequency of the MTF curves cannot be accurately compared, partly because the A system MTF was extrapolated (the dashed line portion) from the trend in the curve, partly
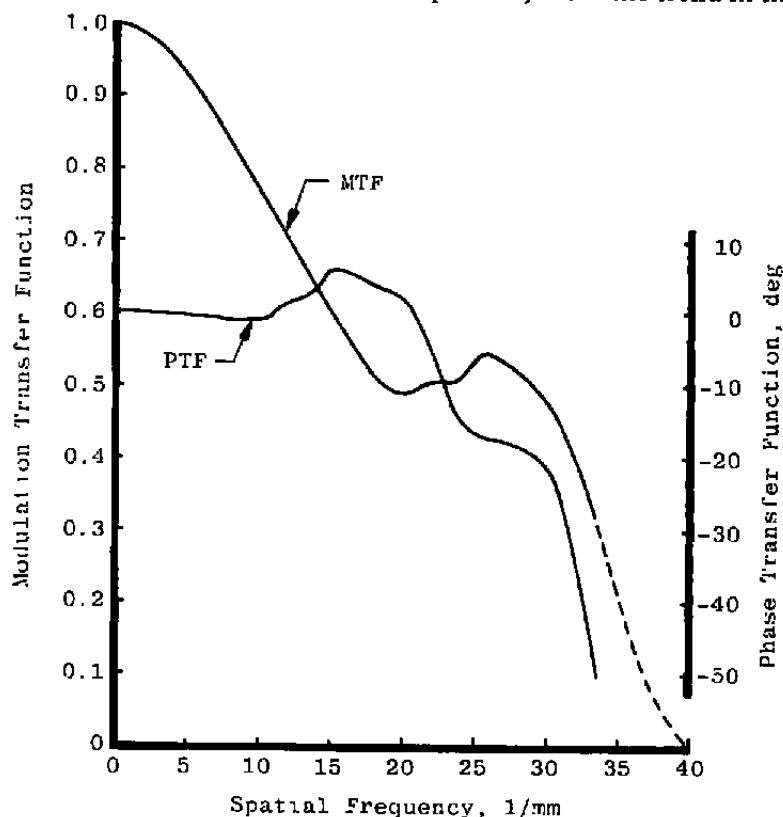

Figure 6. Modulation and phase transfer functions of the A Hologram minus the microdensitometer effects.

because of the measurement uncertainty at higher spatial frequencies. The modulation of the A system is considerably larger past about $f_s = 10 \,(1/\text{mm})$ than that of the B system; the greater area under the MTF curve in the A system loosely implies a better quality of that imaging system (Ref. 18).

Figures 8 and 9 are the experimental line spread functions minus microdensitometer effects of holographic systems A and B, respectively. The shape of the line spread functions in their center portion is as expected from the $\text{sinc}^2 \, [kx_3L/z_2]$ distribution of Eq. (9). The line spread functions are obtained from inverse Fourier transformation of the OTF by numerical
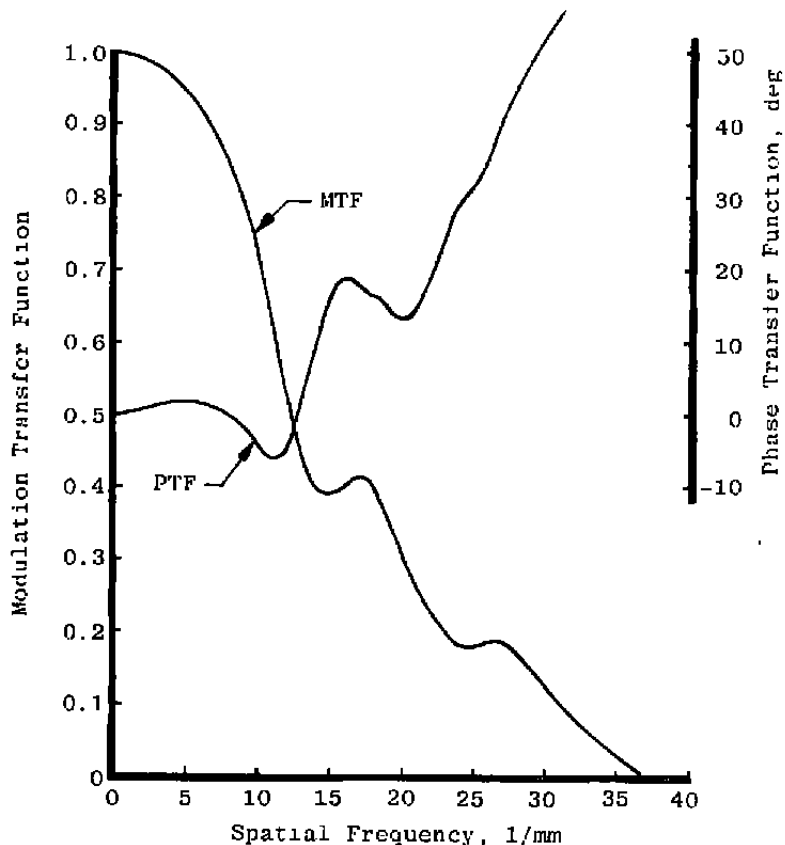
23

Figure 7. Modulation and phase transfer functions of the
B Hologram minus the microdensitometer effects.

integration. Contributions of the higher spatial frequencies of the OTF affect the line spread function width such that erroneously large high-frequency contributions that result from the microdensitometer overcorrections make the line spread functions somewhat too narrow. Thus, Figs. 8 and 9 imply image resolving capabilities of the holographic systems that are perhaps better than their actual capabilities.

6.0 DISCUSSION

The holographic system has incorporated within it the film effects both during the holographic recording and while making a permanent recording of the reconstructed image. These steps introduce nonlinearities because of the nonlinear nature of film recording parameters. While the nonlinearities do not detract greatly from the reconstruction fidelity of the holographic image (Ref. 11), this is not the case when recording that image. The image field consists of the illuminated region which, like the original step function, rather abruptly goes to zero intensity. The greatest demand made on the film is in recording this transition region, yet it is precisely at the transition region that the nonlinearities are most prominent.

Very low levels of illumination in the shadow region of the step function either will be below the threshold of the film sensitivity or will nonetheless be attenuated by the slow threshold response of the film's Huerter-Driffield sensitivity curves. Either way, the effect is that the recorded image shows a somewhat steeper response curve than would be expected of the holographic image (actually, of an image made by any optical system) without the final recording step.

The synergistic effect that increases the slope of the response function is analogous to the case of lens correction through compensating aberrations. This effect also emphasizes one limitation of the transfer function approach to optical analysis. Simply multiplying the component MTF's of a system to obtain the overall system MTF is not sufficiently versatile to describe the image characteristics resulting from compensating effects.

Consequently, the line spread functions of Figs. 8 and 9 might be expected to correlate only qualitatively to the theoretical model expressed by Eq. (9); however, the correlation is
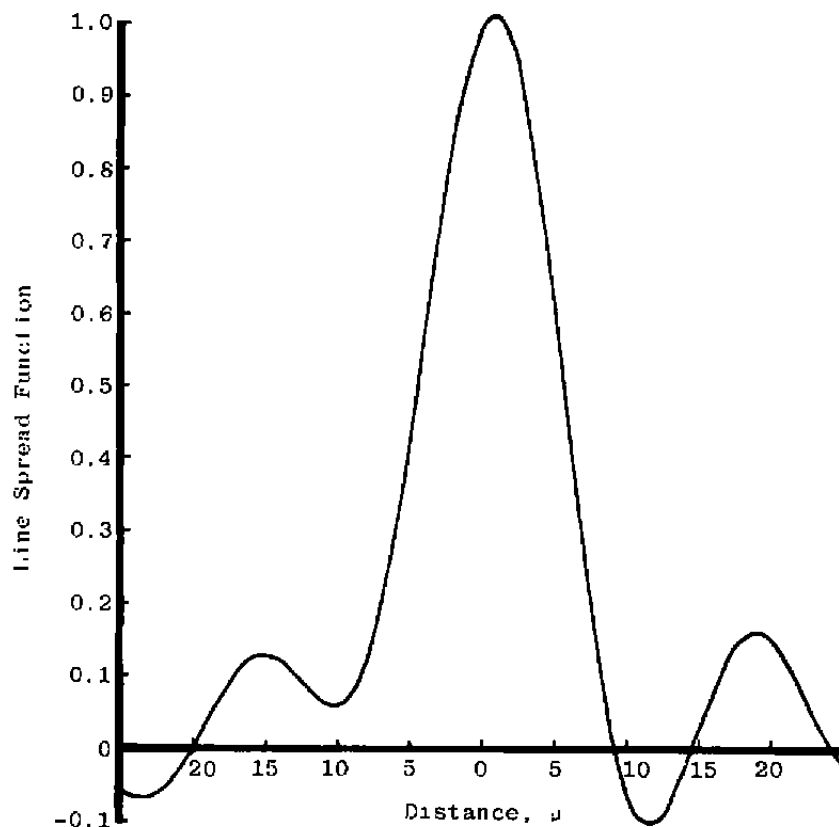

Figure 8. Normalized line spread function of the A hologram
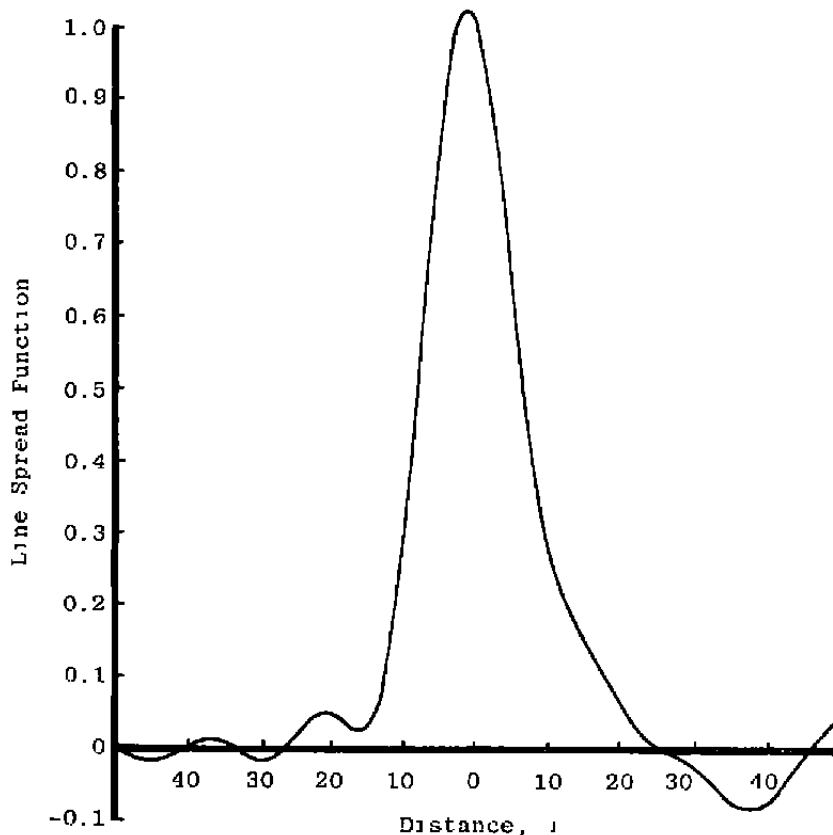minus the microdensitometer effects.

25

Figure 9. Normalized line spread function of the B hologram minus the microdensitometer effects.

quite good. Equation (9) predicts the widths for the parametric conditions of systems A and B, at half-height of the line spread function, to be about 5.6 μm and 17.8 μm, respectively. Figures 8 and 9 show the half widths of about 10 μm and 15 μm as well as the obviously increased spread function of the B system over that of the A system. Application of the Sparrow resolution criterion implies that line groups with line widths and separations of about the same size should be observed. The line groups on a resolution test target observed visually from reconstructions of holograms that were recorded with the same distance and size parameters as holograms A and B could be resolved for line widths of 8 μm and 28 μm, respectively. In these cases the resolution is dependent on subjective interpretation of the images. This, in addition to the resolution restrictions of the microscope-eye combination, again gives a qualitative nature to comparisons with the above results. Nevertheless, the theoretical predictions, the experimentally obtained line spread functions, and the visual observations all gave results that are comparable to within a factor of two.

7.0 SUMMARY

The experimentally obtained line spread function of the real image reconstruction of a holographic system is presented. Two cases which differ in the apertured size of the hologram are shown. The report outlines the theoretical derivations of the holographic reconstruction of point and line spread function and indicates the multiparametric nature of these functions. The application of the spread functions to optical linear systems is discussed, and the linear systems approach is used to detail an experimental technique for obtaining the line spread functions. Although the theoretical and experimental spread functions compare only within limitations (for the reasons given in the discussion section), the resolution of the system implied by both results is nonetheless similar. Furthermore, the shape of the central portion of the experimental line spread functions is as would be expected from the theoretical sinc(x) distribution. Uncertainties in the experimental results that are caused by the experimental procedure's limitations are discussed.

REFERENCES

1. Belz, R. A. and Shofner, F. M. "Characteristic Measurements of an Aperture-Limited In-Line Hologram Image." *Applied Optics*, Vol. 11, No. 10, October 1972, pp. 2215-2222.

2. Carter, W. H. and Dougal, A. A. "Field Range and Resolution in Holography." *Journal of the Optical Society of America*, Vol. 56, No. 12, December 1966, pp. 1754-1759.

3. Champagne, E. B. and Massey, N. G. "Resolution in Holography." *Applied Optics*, Vol. 8, No. 9, September 1969, pp. 1879-1885.

4. Diamond, F. I. "Magnification and Resolution in Wavefront Reconstruction." *Journal of the Optical Society of America*, Vol. 57, No. 4, April 1967, pp. 503-508.

5. Champagne, E. B. "Nonparaxial Imaging, Magnification, and Aberration Properties in Holography." *Journal of the Optical Society of America*, Vol. 57, No. 1, January 1967, pp. 51-55.

6. Meier, R. W. "Magnification and Third-Order Aberrations in Holography." *Journal of the Optical Society of America*, Vol. 55, No. 8, August 1965, pp. 987-992.

7. Meier, R. W. "Depth of Focus and Depth of Field in Holography." *Journal of the Optical Society of America*, Vol. 55, No. 12, December 1965, pp. 1693-1694.

8. Menzel, R. W. "Characteristics of the Holographically Reconstructed Image of a Point Source of Radiation." AEDC-TR-75-131 (AD-A018692), December 1975.

9. Van Ligten, R. F. "Influence of Photographic Film on Wavefront Reconstruction I. Plane Wavefronts." *Journal of the Optical Society of America*, Vol. 48, No. 7, July 1958, pp. 487-490.

10. Friesem, A. A., Kozma, A., and Adams, G. F. "Recording Parameters of Spatially Modulated Coherent Wavefronts." *Applied Optics*, Vol. 6, No. 5, May 1967, pp. 851-856.

11. Yu, F. T. S. *Introduction to Diffraction, Information Processing, and Holography.* MIT Press: Cambridge, 1973.

12. Goodman, J. W. *Introduction to Fourier Optics.* McGraw-Hill: San Francisco, 1968.

13. Born, M. and Wolf, E. *Principles of Optics.* Pergamon Press, Oxford, 1965.

14. Linfoot, E. H. and Wolf, E. "Phase Distribution Near Focus in an Aberration-free Diffraction Image." *Proceedings of the Physics Society*, Vol. 69, Sec. B, No. 8, August 1956, pp. 823-832.

15. Lamberts, R. L., Higgins, G. C., and Wolfe, R. N. "Measurements and Analysis of the Distribution of Energy in Optical Images." *Journal of the Optical Society of America*, Vol. 48, No. 7, July 1958, pp. 487-490.

16. Lamberts, R. L. "Relationship between the Sine-Wave Response and the Distribution of Energy in the Optical Image of a Line." *Journal of the Optical Society of America*, Vol. 48, No. 7, July 1958, pp. 490-495.

17. Perrin, F. H. "Methods of Appraising Photographic Systems. Part I – Historical Review." *Journal of the Society of Motion Picture and Television Engineers*, Vol. 69, No. 3, March 1960, pp. 151-156.

18. Perrin, F. H. "Methods of Appraising Photographic Systems. Part II – Manipulation and Significance of the Sine-Wave Response Function." *Journal of the Society of Motion Picture and Television Engineers*, Vol. 69, No. 4, April 1960, pp. 239-249.

19. Higgins, G. C. and Perrin, F. H. "The Evaluation of Optical Images." *Photographic Science and Engineering*, Vol. 2, No. 2, August 1958, pp. 66-76.

20. Smith, F. D. "Optical Image Evaluation and the Transfer Function." *Applied Optics*, Vol. 2, No. 4, April 1963, pp. 335-350.

21. Scott, F., Scott, R. M., and Shack, R. V. "The Use of Edge Gradients in Determining Modulation-Transfer Functions." *Photographic Science and Engineering*, Vol. 7, No. 6, November-December 1963, pp. 345-349.

22. Shack, R. V. "Characteristics of an Image-Forming System." *Journal of Research of the National Bureau of Standards*, Vol. 56, No. 5, May 1956, pp. 245-260.

23. Jones, R. A. and Coughlin, J. F. "Elimination of Microdensitometer Degradation from Scans of Photographic Images." *Applied Optics*, Vol. 5, No. 9, September 1966, pp. 1411-1414.